REMARKS

Applicant respectfully requests reconsideration and allowance of the subject application. Claims 1-41 are pending.

Typographical Correction

The following paragraphs are amended to correct typographical errors. No new matter has been added.

Fourth paragraph (beginning at line 18) of page 5.

Fifth paragraph (beginning at line 24) of page 5.

Second paragraph (beginning at line 6) of page 19.

First paragraph (beginning at line 1) of page 27.

Claim Objections

Claim 24 is objected to because of the informality of having two duplicate words "with the" on line 8. Claim 24 has been amended to delete a duplicate "with the."

35 U.S.C. §102

Claims 1-3, 11, 27-28 and 31 are rejected under 35 U.S.C. §102 as being anticipated by U.S. Patent No. 5,870,754 to Dimitrova et al (Dimitrova). Applicants respectfully traverse the rejection.

This invention concerns an image retrieval system that performs both keyword-based and content-based image retrieval. A user interface allows a user to specify a query using a combination of keywords and example images. Depending on the input query, the image retrieval system finds images with

keywords that match the keywords in the query and/or images with similar low-level features, such as color, texture, and shape. The system ranks the images and returns them to the user.

The user interface allows the user to identify images that are more relevant to the query, as well as images that are less or not relevant. The image retrieval system monitors the user feedback and uses it to refine any search efforts and to train itself for future search queries.

In the described implementation, the image retrieval system seamlessly integrates feature-based relevance feedback and semantic-based relevance feedback. With feature-based relevance feedback, the system learns which low-level features led to relevant images and groups such features together to aid future searches.

With semantic-based relevance feedback, the image retrieval system learns which keywords are identified with the relevant images and strengthens the associations between the keywords and images. More specifically, the images and keywords are maintained in a database and a semantic network is constructed on top of the image database to define associations between the keywords and images and updated when user feedbacks are provided. Weights are assigned to the keyword-image associations to indicate how relevant the keyword is to the image. The weights are adjusted according to the user feedback, thereby strengthening associations between keywords and images identified as more relevant and weakening the associations between keywords and images identified as less relevant.

Claim 1, for example, recites a method comprising:

initiating a search for images based on at least one query keyword in a query; and

identifying, during the search, first images having associated keywords that match the query keyword and second images that contain low-level features similar to those of the first images.

The method of claim 1 is not disclosed by Dimitrova. Dimitrova shows a method of retrieving video clips using signatures of representative frames (Dimitrova col. 3 lines 1-3). Fig. 12 of Dimitrova shows a display and user interface to request (i.e., query), retrieve, and display video clips to a user. A signature associated with a video clip is represented by the sequence of signatures extracted from representative frames within the video clip (Dimitrova col. 3 lines 44-48). Signatures of video clips are extracted from the video clips and stored in a database (Dimitrova col. 5 lines 1-3). In order to allow for signature extraction, the video clip must have encoded with DC coefficients and motion vectors, without quantization, run-length encoding and Huffman coding. In addition ... a video clip must be Huffman decoded, run-length decoded and dequantized, and have DCT coefficients and motion vectors (Dimitrova col. 5 lines 12-22). The signatures described in Dimitrova are not text based keywords that are annotated to a video frame or video clip (i.e., image). An object of Dimitrova is to retrieve video sequences (i.e., images) without relying on text annotations (text based keywords annotated to an image) (Dimitrova col. 2 lines 66-67). The extracted signatures are stored in a meta database. The term meta database is used because the data (signatures) stored therein described other data such as signatures of other video clips (Dimitrova col. 9 lines 19-24). A user interface allows a user to perform a query and a search for video clips. A user selects an operation such as

"Search by Video", a topic such as "News", and a query video clip. A search is then performed on the extracted signatures stored in the meta database (Dimitrova col. 22 lines 10-20).

Claim 1 recites in part "initiating a search for images based on at least one query keyword in a query." Dimitrova does not disclose this aspect. The video frames and video clips described in Dimitrova have signatures associated with them. Such signatures are not keywords. Dimitrova teaches away from the use of textual annotation in retrieving images. The Examiner identifies a topic such as "News" as being a query keyword. However, "News" is not a keyword associated with any particular video frame or video clip. Further, although a query video clip has an associated signature, the search is initiated based on the user selecting the particular video clip and not the signature associated with the clip.

Claim 1 further recites in part "identifying, during the search, first images having associated keywords that match the query keyword and second images that contain low-level features similar to those of the first images." Dimitrova does not disclose this aspect. As discussed a query keyword is not used to perform the search, therefore without a query keyword, images with associated keywords matching the query keyword cannot be identified. Dimitrova does not disclose identifying second images based on low-level features similar to the first images. The Examiner cites col. 9 lines 44-67 of Dimitrova for disclosing "identifying second images that contain low-level features similar to those of the first images." However, described therein is a signature similarity comparison between a signature of a query video clip and signatures of query video clips in the meta database. Dimitrova discloses determining how close signatures of video frames and video clips are to one another, not what low-level features are similar between

images. Dimitrova discloses that video clips and frames are searched and identified exclusively through their signatures.

For these reasons, claim 1 is patentable over Dimitrova. Applicants respectfully request that the §102 rejection of claim 1 be withdrawn.

Dependent claims 2-3, 11 are allowable by virtue of their dependency on base claim 1. For the reasons given above with respect to claim 1, the systems and methods recited in claims 2-3, 11 are neither disclosed nor taught by Dimitrova. Applicants respectfully request that the §102 rejection of claim 2-3, 11 be withdrawn.

Claim 27 recites an image retrieval system comprising:

> a query handler to handle both keyword-based queries having one or more search keywords and content-based queries having one or more low-level features of an image; and
>
> a feature and semantic matcher to identify at least one of (1) first images having keywords that match the search keywords from a keyword-based query, and (2) second images having low-level features similar to the low-level features of a content-based query.

The method of claim 27 is not disclosed by Dimitrova. As discussed in support of claim 1, Dimitrova shows a method of retrieving video clips using signatures, not keywords. Dimitrova shows a display and user interface to request (i.e., query), retrieve, and display video clips to a user. In Dimitrova, a user selects an operation such as "Search by Video" operation, a topic such as "News", and a query video clip. A search is then performed on the extracted signatures that are stored in the meta database.

Dimitrova does not teach or disclose searching of images based on a content-based query. As discussed above, the query described in Dimitrova is based on an operation such as "Search on Video", a topic such as "News", or a particular query video clip. Dimitrova does not disclose that content-based queries may be made. Further, low-level features associated with images are different than signature similarity between images as disclosed in Dimitrova.

Examiner has not addressed where in Dimitrova is suggested or taught the use of "a feature and semantic matcher to identify at least one of (1) first images having keywords that match the search keywords." Dimitrova fails to teach or disclose a feature and semantic matcher. Dimitrova exclusively uses signatures in searching for images (video frames and video frames). As discussed, Dimitrova teaches away from textual annotations (keywords), therefore a semantic matcher is not relevant in to Dimitrova. Dimitrova further does not teach identifying images by their features.

For these reasons, claim 27 is patentable over Dimitrova. Applicants respectfully request that the §102 rejection of claim 27 be withdrawn.

Dependent claims 28 and 31 are allowable by virtue of their dependency on base claim 27. For the reasons given above with respect to claim 27, the systems and methods recited in claims 28 and 31 are neither disclosed nor taught by Dimitrova. Applicants respectfully request that the §102 rejection of claim 28 and 31 be withdrawn.

35 U.S.C. §103

Claims 4-10, 12-23 and 33-41 are rejected under 35 U.S.C. §103(a) as being unpatentable over Dimitrova in view of U.S. Patent 6,369,811 to Graham et al (Graham).

Independent claim 1 recites "initiating a search for images based on at least one query keyword in a query; and identifying, during the search, first images having associated keywords that match the query keyword and second images that contain low-level features similar to those of the first images."

Claims 4-10 depend from and comprise all the elements of claim 1.

The combination of Dimitrova and Graham fails to teach or suggest the method of claim 1. Neither reference teaches "initiating a search for images based on at least one query keyword in a query." Dimitrova describes initiating a search based on an operation, a topic, and query video clip (Dimitrova col. 22 lines 14-16). Video frames and video clips in Dimitrova are associated with a signature. Dimitrova teaches away from textual annotations in retrieving images and relies on signature association to images, not keyword association to images (Dimitrova col. 2 lines 66-67). The search for images (i.e., video frames and video clips) is performed based on the chosen operation, topic, and query video clip. The search then is performed on signatures in a meta database, the images associated with the returned signatures from the search are presented (Dimitrova col. 22 lines 17-20). Dimitrova does not teach or suggest identifying second images that contain low-level features similar to those of the first images. Dimitrova identifies images through signatures, including performing similarity calculations based on how similar a signature of a query video clip (video frame) is with signatures of other video clips (video frames) (Dimitrova col. 9 lines 52-67). Graham is cited for

teaching monitoring user feedback, however, Graham fails to add any teaching regarding initiating a search for images with a keyword. Graham further fails to add any teaching as to identifying second images that contain low-level features similar to the first images.

Dependent claims 4-10 are allowable by virtue of their dependency on base claim 1. Applicants respectfully request that the §103 rejection of claims 4-10 be withdrawn.

Independent claim 12 recites "a method comprising: permitting entry of both keyword-based queries and content-based queries; finding images using both semantic-based image retrieval and low-level feature-based image retrieval; presenting the images to a user so that the user can indicate whether the images are relevant; and conducting semantic-based relevance feedback and low-level feature-based relevance feedback in an integrated fashion."

The combination of Dimitrova and Graham fails to teach or suggest the method of claim 12. As discussed in support of claims 4-10, neither reference teaches "keyword-based queries." The query described in Dimitrova is based on a chosen operation, a topic, and a query video clip. Dimitrova performs searches based on signatures, not keywords. Dimitrova further fails to suggest or teach that the query may be performed based on content. Graham is cited for teaching monitoring user feedback, however, the Examiner has not addressed where in Graham is suggested or taught the use of "semantic-based relevance feedback and low-level feature-based relevance feedback in an integrated fashion." Accordingly, a combination of Dimitrova and Graham fails to teach or suggest the claimed methods. Applicants respectfully requests that the §103 rejection of claim 12 be withdrawn.

Dependent claims 13-15 are allowable by virtue of their dependency on base claim 12. Applicants respectfully request that the §103 rejection of claims 13-15 be withdrawn.

Independent claim 16 recites a "method comprising: associating keywords with images to form keyword-image links; assigning weights to the keyword-image links; presenting a result set of images obtained from an image retrieval search based on a query; receiving feedback from a user as to whether the images in the result set are relevant to the query; and modifying the weights according to the user feedback."

The combination of Dimitrova and Graham fails to teach or suggest the method of claim 16. As discussed, Dimitrova discloses a signature association with an image. Dimitrova relies on a signature extraction from a video clip, that requires the video clip to have encoded with DC coefficients and motion vectors, without quantization, run-length encoding and Huffman coding. In addition ... a video clip must be Huffman decoded, run-length decoded and dequantized, and have DCT coefficients and motion vectors (Dimitrova col. 5 lines 12-22). In Dimitrova the signature is an inherent property that is unique to the video clip, and cannot be modified. Therefore, Dimitrova does not allow the unique signature to be modified, in particular given a weight. Graham is cited for teaching monitoring user feedback, however, fails to add any teaching regarding assigning weights to the keyword-image links. Accordingly, a combination of Dimitrova and Graham fails to teach or suggest the claimed methods. Applicants respectfully request that the §103 rejection of claim 16 be withdrawn.

Dependent claims 17-19 are allowable by virtue of their dependency on base claim 16. Applicants respectfully request that the §103 rejection of claims 17-19 be withdrawn.

Independent claim 20 recites a "method comprising: presenting a result set of images that are returned from an image retrieval search of a query having at least one keyword; monitoring feedback from a user as to whether the images in the result set are relevant to the query; in an event that the user selects at least one image as being relevant to the query, associating the keyword in the query with the selected image to form a first keyword-image association and assigning a comparatively large weight to the first keyword-image association; and in an event that the user identifies an example image for refinement of the search, associating the keyword in the query with the example image to form a second keyword-image association and assigning a comparatively small weight to the second keyword-image association."

The combination of Dimitrova and Graham fails to teach or suggest the method of claim 20. As discussed in support of claim 16, Dimitrova would not allow a signature associated with an image to be modified, in particular giving a weight to the signature. As discussed Graham is cited for teaching monitoring user feedback, however, fails to add any teaching regarding assigning weights to the keyword-image links. Accordingly, a combination of Dimitrova and Graham fails to teach or suggest the claimed methods. Applicants respectfully request that the §103 rejection of claim 20 be withdrawn.

Dependent claims 21-23 are allowable by virtue of their dependency on base claim 20. Applicants respectfully request that the §103 rejection of claims 21-23 be withdrawn.

Independent claim 27 recites "an image retrieval system comprising: a query handler to handle both keyword-based queries having one or more search keywords and content-based queries having one or more low-level features of an image; and a feature and semantic matcher to identify at least one of (1) first images having keywords that match the search keywords from a keyword-based query, and (2) second images having low-level features similar to the low-level features of a content-based query."

Claims 33-37 depend from and comprise the elements of claim 27.

The combination of Dimitrova and Graham fails to teach or suggest the method of claim 27. As discussed above, Dimitrova associates signatures to video clips, where the signatures are derived from the encoded video clips. Dimitrova teaches the elimination of textual keywords that are annotated (linked) to particular video clips or video frames (i.e., images). Examiner has not addressed where in Dimitrova is suggested or taught the use of "a feature and semantic matcher to identify at least one of (1) first images having keywords that match the search keywords." Dimitrova describes searching of signatures in a meta database based on archiving the signatures, and extracting like signatures based on inherent DC components as derived from the encoded video clips and video frames (Dimitrova Figs. 1 and 2). Graham is cited for teaching monitoring user feedback, however, fails to add any teaching regarding a semantic matcher. Accordingly, a combination of Dimitrova and Graham fails to teach or suggest the claimed methods.

Dependent claims 33-37 are allowable by virtue of their dependency on base claim 27. Applicants respectfully request that the §103 rejection of claims 33-37 be withdrawn.

Independent claim 38 recites a "database structure stored on one or more computer-readable media comprising: multiple image files; multiple keywords; and a semantic network to associate the keywords with the image files, the semantic network defining individual keyword-image links that associate a particular keyword with a particular image file, each keyword-image link having a weight indicative of how relevant the particular keyword is to the particular image file."

The combination of Dimitrova and Graham fails to teach or suggest the database structure of claim 38. As discussed above in support of claims 33-37, the Examiner has not addressed how Dimitrova discloses the use of a semantic matcher. As to claim 38, the Examiner has not addressed where in Dimitrova is disclosed the use of a semantic network. As discussed in support of claims 33-37, the use of semantics relates to text, and is different than the search technique based on derived signatures as described in Dimitrova. Graham is cited for teaching monitoring user feedback, however, fails to add any teaching regarding a semantic matcher. Accordingly, a combination of Dimitrova and Graham fails to teach or suggest the claimed methods. Applicants respectfully request that the §103 rejection of claim 38 be withdrawn.

Independent claim 39 recites "a computer-readable medium having computer-executable instructions that, when executed, direct a computer to: find images using both semantic-based image retrieval and low-level feature-based image retrieval; present the images to a user so that the user can indicate whether the images are relevant; and concurrently conduct semantic-based relevance feedback and low-level feature-based relevance feedback.

The combination of Dimitrova and Graham fails to teach or suggest the computer-readable medium of claim 39. As discussed in support of claims 33-38, the use of semantics relates to text, and is different than the search technique based on derived signatures as described in Dimitrova. Further, as discussed above, Dimitrova does not suggest or teach the use of low-level featured-based retrieval, since Dimitrova relies exclusively in the use of derived signatures. Graham is cited for teaching monitoring user feedback, however, fails to add any teaching regarding a semantic matcher. Accordingly, a combination of Dimitrova and Graham fails to teach or suggest the claimed methods. Applicants respectfully request that the §103 rejection of claim 39 be withdrawn.

Dependent claims 40-41 are allowable by virtue of their dependency on base claim 39. Applicants respectfully request that the §103 rejection of claims 40-41 be withdrawn.

Claims 24-26 are rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent 6,285,995 to Abdel-Mottaleb et al (Abdel-Mottaleb) in view of U.S. Patent 5,594,809 to Kopec et al (Kopec).

Amended independent claim 24 recites "a method comprising: computing, for each category, a representative feature vectors of a set of existing images within the category; determining a set of representative keywords that are associated with the existing images in each category; comparing, for each new image, the low-level feature vectors of the new image to the representative feature vectors of the existing images in each category to identify a closest matching category; and labeling the new image with the set of representative keywords associated with the closest matching category."

The combination of Abdel-Mottaleb and Kopec fails to teach or suggest the method of claim 29. Abdel-Mottaleb describes representing the content of an image with a feature vector, and the images are grouped in clusters (Abdel-Mottaleb col. 5 lines 15-19). Images in a database are grouped by clusters (Abdel-Mottaleb col. 5 lines 12-14). A query image and a feature vector for the query image are entered. A comparison is made as to the query image and clusters in the database. For each of the clusters, a cluster similarity value is calculated based on a *cluster center* of each cluster and the respective feature vector of the query image. For the query image, a selection unit selects clusters with the highest cluster similarity value (Abdel-Mottaleb col. 5 lines 35-56). The *cluster center* may either be the average of the color histograms of a number of representative images in the cluster or a feature vector of a single image chosen as the representative image for all images in the cluster (Abdel-Mottaleb col. 5 lines 30-34). Abdel-Mottaleb teaches away from the use of keywords that are annotated to images in performing the search (Abdel-Mottaleb col. 5 lines 8-12). In other words keywords are not attached to images in Abdel-Mottaleb. Therefore Abdel-Mottaleb fails to suggest or teach "determining representative keywords associated with the existing images in each category (i.e., cluster)." In addition, Abdel-Mottaleb fails to suggest or teach "comparing, for each new image, the low-level feature vectors of the new image to the representative feature vectors of the existing images in each category to identify a closest matching category." Abdel-Mottaleb describes calculating a cluster similarity value that uses a cluster center which is an average of the color histograms or is a single feature vector of a representative image. Kopec is cited for "labeling the new image with the set of representative keywords with the closest matching category." Kopec describes

segmenting an input keyword image into image segments, and providing labels for each image segments (Kopec col. 7, lines 63-65). Kopec describes labeling image segments, however does not disclose labeling such image segments with a set of representative keywords that are associated with a category. Accordingly, a combination of Abdel-Mottaleb and Kopec is improper. Applicant respectfully requests that the §103 rejection of claim 24 be withdrawn.

Dependent claims 25-26 are allowable by virtue of their dependency on base claim 24. Applicants respectfully request that the §103 rejection of claims 25-26 be withdrawn.

CONCLUSION

All pending claims 1-41 are in condition for allowance. Applicant respectfully requests reconsideration and prompt issuance of the subject application. If any issues remain that prevent issuance of this application, the Examiner is urged to contact the undersigned attorney before issuing a subsequent Action.

Respectfully Submitted,

Dated: 2/10/03

By: _____

Emmanuel A. Rivera
Reg. No. 45,760
(509) 324-9256 ext. 245

MARKED UP VERSION OF PENDING CLAIMS UNDER 37 C.F.R. § 1.121(C)(1)(ii):

Amend claim 24 as follows and in accordance with 37 C.F.R. § 1.121(c)(1)(ii), by which the Applicant submits the following marked up version only for claims being changed by the current amendment, wherein the markings are shown by brackets (for deleted matter) and/or underlining (for added matter):

24. (Once Amended) A method comprising:

computing, for each category, a representative feature vectors of a set of existing images within the category;

determining a set of representative keywords that are associated with the existing images in each category;

comparing, for each new image, the low-level feature vectors of the new image to the representative feature vectors of the existing images in each category to identify a closest matching category; and

labeling the new image with the [with the]set of representative keywords associated with the closest matching category.

MARKED UP VERSION OF SPECIFICATION UNDER 37 C.F.R. § 1.121(B)(1)(iii):

Page 5, fourth paragraph

An image retrieval system performs both keyword-based and content-based image retrieval. A user interface allows a user to specify a query using a combination of keywords and example[s] images. Depending on the input query, the image retrieval system finds images with keywords that match the keywords in

the query and/or images with similar low-level features, such as color, texture, and shape. The system ranks the images and returns them to the user.

Page 5, fifth paragraph

The user interface allows the user to identify images that are more relevant to the query, as well as images that are less or not relevant.[.] The image retrieval system monitors the user feedback and uses it to refine any search efforts and to train itself for future search queries.

Page 19, second paragraph

Fig. 6 shows an example of a query screen 600 presented by the user interface 200 for entry of an initial query. The screen display 600 presents a natural language text entry area 602 that allows user to enter keywords or phrases. After entering one or more keywords, the user actuates a button 604 that initiates the search for relevant images. Alternatively, the user can browse a pre-defined concept hierarchy by selecting one of the categories listed in section 606 of the query screen 600. The user actuates the category link to initiate a search for images within the category.

Page 27, first paragraph

Because the low-level features are not enough to present the images' semantics, some or even all of the automatically labeled keywords will inevitably be inaccurate. However, through user queries and feedbacks, semantically accurate keywords [labels]will emerge while semantically inaccurate keywords will slowly be eliminated.